NPSOR-91-06

NAVAL POSTGRADUATE SCHOOL

Monterey, California

MAXIMIZATION ON MATROIDS WITH RANDOM WEIGHTS

Michael P. Bailey

January 1991

Approved for public release; distribution is unlimited

Prepared for:
USMC Warfighting Center, MCCDC
Quantico, VA 22134-5001

NAVAL POSTGRADUATE SCHOOL,
MONTEREY, CALIFORNIA

This report was sponsored and funded by the USMC Warfighting Center, MCCDC, Quantico, VA 22134-5001.

This report was prepared by:

MICHAEL P. BAILEY
Asst. Prof. of Operations Research

Reviewed by:

PETER PURDUE
Professor and Chairman
Department of Operations Research

Released by:

PAUL J. MARTO
Dean of Research

Unclassified

Security Classification of this page

REPORT DOCUMENTATION PAGE

1a Report Security Classification UNCLASSIFIED	1b Restrictive Markings
2a Security Classification Authority	3 Distribution Availability of Report Approved for public release; distribution is unlimited
2b Declassification/Downgrading Schedule	
4 Performing Organization Report Number(s) NPSOR-91-06	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School 6b Office Symbol *(If Applicable)* OR	7a Name of Monitoring Organization
6c Address *(city, state, and ZIP code)* Monterey, CA 93943-5000	7b Address *(city, state, and ZIP code)*
8a Name of Funding/Sponsoring Organization USMC Warfighting Center, MCCDC 8b Office Symbol *(If Applicable)* OR/Ba	9 Procurement Instrument Identification Number M9545091WRR1AK2
8c Address *(city, state, and ZIP code)* Quantico, VA 22134-5001	10 Source of Funding Numbers

Program Element Number	Project No	Task No	Work Unit Accession No

11 Title *(Include Security Classification)* MAXIMIZATION ON MATROIDS WITH RANDOM WEIGHTS

12 Personal Author(s) Bailey, Michael P.

13a Type of Report	13b Time Covered	14 Date of Report *(year, month,day)*	15 Page Count
Technical	From To	1991, January	

16 Supplementary Notation The views expressed in this paper are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes

Field	Group	Subgroup

18 Subject Terms *(continue on reverse if necessary and identify by block number)* Matroid dual; Stochastic matroid maximization; Stochastic spanning tree; Stochastic scheduling

19 Abstract *(continue on reverse if necessary and identify by block number*

In this work we develop a method for analyzing maximum weight selections in matroids with random element weights, especially exponentially distributed weights. We use the structure of the matroid dual to transform matroid maximization into an equivalent minimization task. We model sample paths of the greedy minimization scheme using a Markov process, and thus solve the original maximization problem. The distribution of the weight of the optimal basic element and moments are found, as well as the probability that a given basic element is optimal. We also derive criticality indices for each ground set element, giving the probability that an element is a member of the optimal solution. We give examples using spanning trees and scheduling problems, each example being a new result in stochastic combinatorial optimization.

20 Distribution/Availability of Abstract [X] unclassified/unlimited [ ] same as report [ ] DTIC users	21 Abstract Security Classification Unclassified
22a Name of Responsible Individual M. P. Bailey	22b Telephone *(Include Area code)* (408) 646-2085 — 22c Office Symbol OR/Ba

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted — security classification of this page

All other editions are obsolete — Unclassified

A-1

1. INTRODUCTION

There exist several works in the literature on the performance measures of minimum weight spanning trees in networks with randomly weighted arcs. Among these are Frieze [5], and Jain and Mamer [6]. Each of these papers relies on identically distributed arcs, special graph structures, asymptotic exponentiality of lower order statistics, and the greedy minimizing algorithm to characterize performance trends as the size of the network grows. In Kulkarni [7], a method was developed for producing exact probability distributions of the *minimum* weight spanning tree's weight when the network has general structure with nonidentical, exponentially distributed weights. Kulkarni constructs a Markov process whose sample paths are the probable execution paths of Krushal's minimum weight spanning tree algorithm. Several useful by-products emerge from this method, including methods for deriving moments, criticality indices, and conditional distributions. In [2], this method is extended to general matroids with exponentially distributed element weights. The problem studied was to find the distribution of the weight of the minimum weight basic element in a matroid, as well as criticality indices and moments.

In this work, we endeavor to extend the Markov process method to solve matroid maximization problems. Although matroid theory usually addresses maximization problems, there exists only a small amount of literature concerning maximization problems with random element weights, and none concerning analytic methods for matroid maximization. Applications of this method will include maximum weight spanning trees, maximum weight transversals useful in assignment problems (solutions *and* heuristics), optimal scheduling, and flow network synthesis.

2. MATROID COMBINATORICS

In this section we concisely describe the combinatorial underpinnings required to explore matroid maximization. A less terse treatment of the subject of matroids and their applications may be found in Lawler [8]. This section will cover matroid greedy algorithms, matroid structures, and dual matroids.

Let E be a set of objects such as vectors, nodes, or arcs. $\mathcal{M}$ is a set of subsets of F with the following two properties:

2.1) $Y \in \mathcal{M}$ and $X \subset Y$, then $X \in \mathcal{M}$;

2.2) for every subset A of E, $\{X \subset A$: $X \in \mathcal{M}$: there exists no $x \in A$ such that $X \cup \{x\} \in \mathcal{M}\}$ is a set of equicardinal sets.

$\mathcal{M}$ is called the set of ***feasible*** elements. Property 2.1 says every subset of a set in $\mathcal{M}$ is in $\mathcal{M}$. Thus $\mathcal{M}$ is called ***simplical***. Property 2.2 dictates that for every $A \subset E$, every maximal feasible subset of a set A contains the same number of elements. We will denote the set of maximal elements in $\mathcal{M}$ as $\beta_{\mathcal{M}}$, called the ***basis*** of $\mathcal{M}$ and will call members of $\beta_{\mathcal{M}}$ ***basic*** elements. We will consistently use n to denote the cardinality of a basic element.

Properties 2.1 and 2.2 combine to guarantee that we can begin with the empty set $\varnothing$, and construct any set in $\beta_{\mathcal{M}}$ by making n selections from the set E. We will perform this construction of a basic element by greedy minimization.

Let v be a nonnegative weight function on the set E, v: $E \rightarrow \Re^{+}$. The linear objective function ω on elements of $\mathcal{M}$ is given by

$$\omega(x) = \sum_{x \in X} v(x)$$

for each $X \in \mathcal{M}$. The notion of greediness is formalized by the following algorithm:

0. initialize: $X = \varnothing$, $\omega = 0$
1. $x \leftarrow arg\ max_{y:\ X \cup y \in \mathcal{M}}\ v(y)$
2. $w \leftarrow w + v(x)$
3. $X \leftarrow X \cup x$
4. if $X \notin \beta_{\mathcal{M}}$ then go to step 1
5. stop

Figure 1. The Greedy Algorithm

Verbally, the greedy algorithm begins with the empty set $X = \varnothing$ and at each stage selects the element $x \in E - X$ with smallest weight, subject to the constraint that $X \cup \{x\}$ is in $\mathcal{M}$. Let X^G be the basic element constructed by the greedy algorithm, $X^G = \{x_1^G, x_2^G, \ldots, x_n^G\}$ where x_i^G is the element selected at the i^{th} opportunity. Let $X_i^G = \{x_1^G, x_2^G, \ldots, x_i^G\}$ be the set of the first i greedy selections. Note that the terminal value of ω is equal to $\omega(X^G)$, the linear objective function value of X^G. The critical connection between the greedy algorithm and matroid structures is given by the following theorem, directly adapted from results of Edmonds [3].

Theorem 1. Let ω be a linear objective function for an arbitrary nonnegative weight function v, then $X^G = arg\ min_{Y \in \beta_{\mathcal{M}}} \omega(Y)$ if and only if $\mathcal{M}$ is a matroid.

Proof. See Lawler [8]. •

The above theorem holds for finding the minimum weight basic element as well, where "arg min" replaces "arg max" in the greedy algorithm.

In [2], the matriod minimization problem was investigated. A method was given for finding the joint distribution $\{F_X(t),\ X \in \beta_{\mathcal{M}}, t \geq 0\}$, where $F_X(t) = P[X^G = X, \omega(X^G) \leq t]$, when $\{v(x):\ x \in E\}$ is a set of independent, exponentially distributed random weights. This was achieved by constructing a Markov process Z with absorption time distributions $F_X(t) = P[Z(t) = X]$ for each $X \in \beta_{\mathcal{M}}$. We will now use results concerning dual matroids to produce similar results for matroid maximization problems.

Theorem 2. (Minty) $\beta_{\mathcal{M}_c} = \{X_c = E - X: X \in \beta_{\mathcal{M}}\}$ in the basis of a matroid. This complementary matroid is called the dual matroid of $\mathcal{M}$.

Proof. See Lawler [8, p.277] •

Corollary 1. $X_c^G = E\text{-}X^G$ is the maximum weight basic element in $\beta_{\mathcal{M}_c}$.

Proof. Let $T = \sum_{x \in E} v(x)$, then $\omega\left(X_c^G\right) = T\text{-}\omega(X^G) \geq T\text{-}\omega(X) = \omega(X_c)$ for all $X \in \beta_{\mathcal{M}}$. •

Thus, by identifying the minimum weight basic element in $\mathcal{M}$, we may find the maximum weight basic element in the dual matroid $\mathcal{M}_c$. In the sequel, we will consider the case where the set $\{v(x): x \in E\}$ is a set of independent, exponentially distributed random variables.

To explain our restructured objective function, we must discuss the concept of rank. Let $S \subset E$. The cardinality of a maximal feasible subset of S is called the ***rank*** of S, denoted $r(S)$,

$$r(S) = \max_{X \subset S,\, X \in \mathcal{M}} |X|$$

Restating property 2.2, for any set $S \subset E$, $\{X \subset S: X \in \mathcal{M} \text{ and } |X| = r(S)\}$ is the basis of a matroid, the one generated by contracting $\mathcal{M}$ by $E - S$. Thus, $r(E) = n$. Let $n_c = |E| - n$, and note that n_c is the cardinality of a basic element in $\mathcal{M}_c$, or the rank of E with respect to $\mathcal{M}_c$.

We will consider the case where the weights of the ground set elements are random variables with known distributions. Our goal is to model the behavior of the greedy algorithm as a stochastic process. Previous attempts have been thwarted by the complexity of the conditioning arguements required to do this. Every greedy selection yields information about the elements not selected as well as the selection made, so that analyzing the algorithm in its current state would involve managing all of the information concerning the residual costs of nonselected elements at each stage. We will not face this problem directly, choosing instead to reformulate the matroid maximization problem by restructuring the objective function ω.

Our new objective function, denoted ω_D and named the **dual discounted objective function**, is defined on all subsets of E;

$$\omega_D(S) = \max_{X \subset S, X \in \mathcal{M}, |X| = r(S)} \sum_{x \in S - X} v(x) + \left\{\left[n_c - (|S| - r(S))\right]\left[\max_{x \in S} v(x)\right]\right\}.$$

This objective function, though rather unusual in appearance, comes directly from performing the straightforward algorithm shown in figure 2, called the dual greedy algorithm with discounting.

0. initialize: $X = X_c = S = \varnothing$, $\omega = 0$, $u(x) = v(x)$ for each $x \in E$
1. $x \leftarrow arg\ min_{x' \in E\text{-}S}\ u(x')$
2. $S \leftarrow S \cup \{x\}$
3. if $X \cup \{x\} \in \mathcal{M}$
   $X \leftarrow X \cup \{x\}$
   else
   $X_c \leftarrow X_c \cup \{x\}$
4. $w \leftarrow w + \left[n_c - (|S| - r(S))\right]u(x)$
5. $u(y) \leftarrow u(y) - u(x)$ for all $y \in E - S$
6. if $S \neq E$, go to step 1.
7. stop

Figure 2. The Dual Greedy Algorithm with Discounting

Verbally, the dual greedy algorithm with discounting starts with three empty sets, X, X_c, and S. One by one, x, the minimum weight element of E - S is determined. If adding x to X is feasible with respect to $\mathcal{M}$, it is added. Otherwise, x is added to the complementary set X_c. In either case, x is added to the set S, and all the elements remaining in E - S are discounted further by the current discounted value of x. The remainder of this section is dedicated to showing that, at termination of the above algorithm, $\omega = max_{Y \in \beta_{\mathcal{M}_c}} \omega(Y)$ and that X_c is the maximum weight basic element of $\mathcal{M}_c$.

Lemma 1. Given sets S, X, and X_c constructed of some stage of the dual greedy algorithm with discounting,

$$\left| X_c^G - X_c \right| = \left[n_c - (|S| - r(S)) \right]$$

Proof. By property 2.2, we know that $|X| = r(S)$ at every stage of the algorithm. Thus, $|X_c| = |S| - r(S)$ always, while $\left|X_c^G\right| = n_c$ because $\mathcal{M}_c$ is a matroid, hence it has equicardinal basic elements. Because $X_c^G = E\text{-}X^G$ and $X \subset X^G$, $X_c \subset X_c^G$ and the lemma follows. •

Thus, at each stage of the dual greedy algorithm with discounting, we claim that X_c is a subset of the maximum weight basic element of $\mathcal{M}_c$, and that there are exactly $\left[n_c - (|S| - r(S)) \right]$ elements in the set $E\text{-}S$ which we will add to X_c to form X_c^G. Each of these remaining elements has value of at least $max_{y \in S}\, v(y)$. Thus each element of X_c^G has been discounted by $max_{y \in S}\, v(y)$, so the sum of these discounts is

$$\left[n_c - (|S| - r(S)) \right] max_{y \in S}\, v(y).$$

Example 1. Understanding the operation of the dual greedy algorithm with discounting is facilitated by the diagram in Figure 3, illustrating the operation of the discounted greedy algorithm for the matroid with basis $\beta_{\mathcal{M}} = \{\{x_2, x_4\}, \{x_1, x_4\}, \{x_4, x_5\}, \{x_3, x_5\}\}$, thus $\beta_{\mathcal{M}_c}$ $\{\{x_1, x_3, x_5\}, \{x_2, x_3, x_5\}, \{x_1, x_2, x_3\}, \{x_1, x_2, x_4\}\}$ $n = 2$, $n_c = 3$. The values of the weight function v are given in the first line of table 1. By inspection, $X^G = \{x_1, x_4\}$ and thus $\{x_2, x_3, x_5\} = X_c^G$. The shaded area in the figure indicates the magnitude of the value of $\omega(\{x_1, x_2\})$ which is 21. This value is the sum of $v(x_2)$ plus the discount taken from the $v(x_3)$ and $v(x_5)$, which we will eventually absorb. Table

1 shows the execution path of the dual greedy algorithm with discounting on $\mathcal{M}$ with the weights specified.

	$v(x_1) = 6$
	$v(x_2) = 7$
	$v(x_3) = 9$
	$v(x_4) = 12$
	$v(x_5) = 16$

Shaded area = $\omega_{\mathcal{c}}(\{x_1, x_2\})$

Figure 3. Accumulating Cost in the Dual Greedy Algorithm with Discounting

TABLE 1. THE EXECUTION OF DUAL GREEDY ALGORITHM WITH DISCOUNTING

ITERATION	S	X	X_c	ω_p	$u(x_1)$	$u(x_2)$	$u(x_3)$	$u(x_4)$	$u(x_5)$	$r(S)$	$rc(E \setminus S) - r(S)$
0	$\emptyset$	$\emptyset$	$\emptyset$	0	6	7	9	12	16	0	3
1	$\{x_1\}$	$\{x_1\}$	$\emptyset$	18	6	7	9	12	16	1	3
2	$\{x_1, x_2\}$	$\{x_1\}$	$\{x_2\}$	20		1	3	6	10	1	2
3	$\{x_1, x_2, x_3\}$	$\{x_1\}$	$\{x_2, x_3\}$	23	-	-	2	5	9	1	1
4	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_4\}$	$\{x_2, x_3\}$	25		-	-	3	7	2	1
5	$\{x_1, x_2, x_3, x_4, x_5\}$	$\{x_1, x_4\}$	$\{x_2, x_3, x_5\}$	32	-	-	-	-	4	2	0

Theorem 3. Let the dual greedy algorithm with discounting terminate with sets X, X_c, and the value ω. Then

i) $X = X^G$ = the maximum weight basic element in $\mathcal{M}$.

ii) $X_c = X_c^G$ = the maximum weight basic element in $\mathcal{M}_c$.

iii) $\omega = \sum_{x \in X_c^G} v(X_c)$.

Proof. Statement i is a direct result of theorem 1. Statement ii is true by using statement i in conjunction with the corollary to Minty's theorem on matroid duals. Statement iii relies on lemma 1. •

3. MATROID MAXIMIZATION WITH EXPONENTIALLY DISTRIBUTED ELEMENT WEIGHTS

The motivation for developing the dual greedy algorithm with discounting is that, given that the element weights are independent, exponentially distributed random variables, the execution paths of the algorithm are the sample paths of a Markov process. Let $\{V(X): X \in E\}$ be a set of independent, exponentially distributed weights with rates $\{\lambda(x): x \in E\}$, and W_D be the associated stochastic dual discounted objective function.

Let Z be a Markov process with state space $2^{|E|}$, initial state $\varnothing$, absorbing state E, and transition rates given by

$$Q_{S,S\cup\{x\}} = \frac{\lambda(x)}{n_c - (|S| - r(S))}. \tag{3.1}$$

for any set $S \neq E$. The elements of the lower half of Q are all zero, and the diagonal elements are always -1 times the sum of the elements in the row.

Theorem 4.

$$P[Z(t) = E] = P\left[W_D\left(X_c^G\right) \le t\right].$$

Proof. By the construction of the dual greedy algorithm with discounting, we have that for any set of realizations of $\{V(x): X \in E\}$ and at any stage of the algorithm, the increase in the value of W_D upon the next iteration is given by

$$\left[n_c - (|S| - r(S))\right] \min_{y \in E-S} u(y)$$

At the outset, $u(y) = V(y)$ for each $y \in E$. As elements are added to the set S, the operation $u(y) \leftarrow u(y) - u(x)$ is performed, where $u(x) = \min_{y \in E-S} u(y)$. Thus for $y \notin S$ after iteration k, $u(y)$ is the residual value of y given that $V(y)$ is not one of the k smallest members of $\{v(y') : y' \in E\}$. Thus, by the strong Markov property, each $y \in S$ has $u(y) \sim exp(\lambda(y))$. Hence, given transition times $\tau_1, \tau_2, \ldots, \tau_i$,

$$P\left[Z(\tau_{i+1}) = S \cup \{x\} \middle| Z(\tau_i) = S\right]$$

$$= P\left[u(x) \le u(y) \text{ for all } y \in E - S, \left[n_c - (|S| - r(S))\right] u(x) \le \tau_{i+1} - \tau_i\right]$$

$$= \frac{\lambda(x)}{\sum_{y \in E-S} \lambda(y)} \cdot \exp\left[-\sum_{y \in E-S} \frac{\lambda(y)}{\left[n_c - (|S| - r(S))\right]}\right].$$

$$= \frac{Q_{S, S \cup \{x\}}}{-Q_{S,S}} e^{Q_{S,S}}$$

Hence, the sojourn time in each state is exponentially distributed, and the Markov process with rate matrix Q has the property that its absorption time and $W_D\left(X_c^G\right)$ are identically distributed. •

In the suceeding sections, we will apply this theorem to two important matroids.

We can derive the distribution of $W_D\left(X_c^G\right)$ as the time of absorption of the Markov process Z. The remaining issue to resolve is which element of $\beta_{\mathcal{M}_c}$ is optimal. It is clear that the sample path of Z will reveal the optimal element of $\beta_{\mathcal{M}_c}$, it is the element whose complement is contained in S at the earliest stage. We will determine this element using the embedded Markov chain of Z.

Truncate the statespace of Z so that each state containing an element in $\beta_{\mathcal{M}}$ is absorbing, calling the new process Z'.

Corollary 2: For any $S \subset E$ containing an element X of $\beta_{\mathcal{M}}$,

$$P[Z' \text{ is absorbed in } S]$$
$$= P[X = X_G]$$
$$= P[X_c = X_c^G].$$

The astute reader will realize that implementing the method suggested in corollary 2 is quite wasteful. What is truly being identified by Z' is the minimum weight element in the primal matroid $\mathcal{M}$. This is the focus of the companion work, [2]. We draw the connection in the following corollary.

Corollary 3. Let $\{Y_j, j \geq 0\}$ be a Markov chain with $Y_0 = \varnothing$ and transition probability matrix given by

$$P_{X, X \cup \{x\}} = \frac{\lambda(x)}{\sum_{X \cup \{y\} \in \mathcal{M}} \lambda(y)} \text{ for } X, X \cup \{x\} \in \mathcal{M}.$$

with the elements of the lower half of P are all zero, as are all of the diagonal elements. Then $P[Y_n = X] = P\left[X_c = X_c^G\right]$.

Proof. There exists a Markov process Z'' on statespace $\mathcal{M}$ such that $Z''(0) = \varnothing$ and Z'' is controlled by the generator Q'' given by

$$Q''_{X, X \cup \{x\}} = \frac{\lambda(x)}{(n - |X|)}.$$

for $X \cup \{x\}$ in $\mathcal{M}$. This process has the property that $P[Z''(t) = X \in \beta_{\mathcal{M}}] = P[X_G = X, W(X^G) \leq t]$. This is the fundamental result of [2], and is provable using techniques very similar to those used in theorem 3.1. Note that $|X| = r(X)$ due to the restricted statespace of Z''. The embedded Markov chain of Z'', which we denote by Y, has transition probability matrix given by

$$P_{X, X \cup \{x\}} = \frac{(n - |X|)\lambda(x)}{\sum_{y \in X \cup \{y\} \in \mathcal{M}} \lambda(y)(n - |X|)} = \frac{\lambda(x)}{\sum_{y : X \cup \{y\} \in \mathcal{M}} \lambda(y)}.$$

Thus,

$$\lim_{t \to \infty} P[Z''(t) = X \in \beta_{\mathcal{M}}] = P[Y_n = X] = P\left[X_c = X_c^G\right].$$

Corollary 4.

$$P\left[x \in X_c^G\right] = P[x \notin Y_n].$$

for all $x \in E$.

This last corollary gives the probability that an element of E is a member of the maximum weight basic element, a probability often referred to as the ***criticality index*** of the element in question. Criticality indices can play a major role in allocating resources to improve the performance of a matroid maximization, as

$$P\left[x \in X_c^G\right] = P\left[\frac{\partial}{\partial V(x)} W_D\left(X_c^G\right) \neq 0\right].$$

4. AN EXAMPLE IN SCHEDULING

The second most widespread use of weighted matroids is in the area of production scheduling on a single processor. This problem is solved using the transversal matroid of Edmonds and Fulkerson [4] on a convex bipartite graph.

Suppose that we have N tasks with earliest starting times $b_1, b_2, \ldots, b_N$ and latest permissible completion deadlines $d_1, d_2, \ldots, d_N$. We have a single machine on which the tasks can be performed, each task taking one unit of time. The premium paid for accomplishing job i on time is $v(i)$, no premium is paid if a job is not processed in its permissible time interval. Our goal is to schedule all of the jobs so that the sum of the paid premiums is maximized. We will construct a matroid so that the maximum weight basic element gives the optimal schedule.

Let $T_1 = \{1, 2, \ldots, \max_i(d_i)\}$ and $T_2 = \{1, 2, \ldots, N\}$. Let (T_1, T_2, A) be a bipartite graph where $(t,k) \in A$ if $b_k \leq t \leq d_k$, that is, job k is connected to every time period in which it may be scheduled. A maximum benefit schedule consists of a matching of time periods to tasks, where any task matched to a time period leads to payment of the associated premium. As a matter of interest, the reader should know that bipartite graphs where $\{t: (t,k) \in A\}$ is a set of consecutive integers for each $k \in T_2$ is called **convex**.

The following policy always generates the optimal schedule:

POLICY: Choose the remaining task with the highest premium and attempt to schedule it with the other tasks chosen thus far. Stop when this cannot be done.

Thus $\mathcal{M}_c = \{X \subset T_2$: X matches to some subset of $T_1\}$. For illustration, let us consider six tasks given in Table 2.

TABLE 2. TASKS, STARTING TIMES, DEADLINES, AND PREMIUMS

TASK	BEGIN	DEADLINE	PREMIUM
1	1	1	6
2	1	2	9
3	2	3	12
4	1	2	4
5	2	4	6
6	3	4	8

This set of tasks and deadlines leads to the convex bipartite network given in Figure 4.

The dual greedy algorithm with discounting will follow the execution path $\varnothing$, $\{4\}$, then either $\{1,4\}$, or $\{4,5\}$ (there is a tie), it will then continue until all tasks are scheduled and yield the value $\omega_D\left(X_C^G\right)$ as $\omega_D(\{1,2,3,6\}) = \omega_D(\{2,3,5,6\}) = 35$.

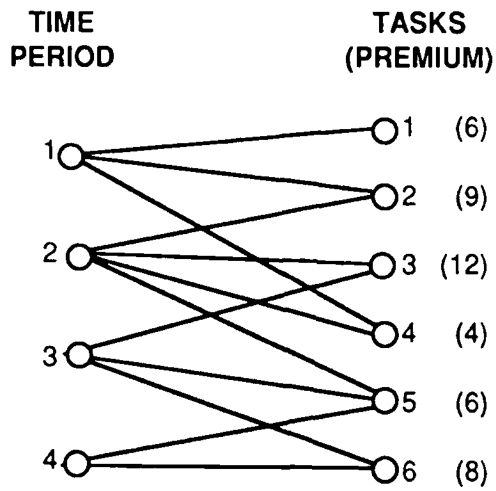

Figure 4. Convex Bipartite Graph Connecting Tasks to Time Periods

Now suppose that each benefit is exponentially distributed with mean equal to the benefit in Table 2. We constructed the Markov process Z as per equation 3.1, and computed performance measures. Figures 5 and 6 give the density and distribution functions, resp., for the random benefit of the optimal schedule.

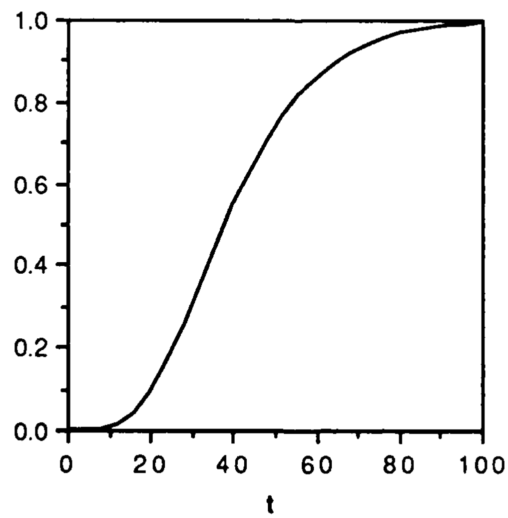

Figure 5. The Distribution of $W_D\left(X_c^G\right)$

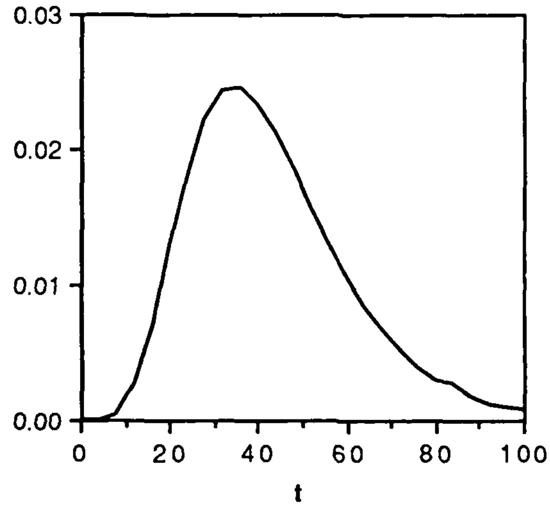

Figure 6. The Density of $W_D\left(X_c^G\right)$

The density shows a mode near 40, with a fairly generous tail to the right. The shape of both the density and the distribution function are typical for small applications.

The moments of $W_D\left(X_C^G\right)$ may be found using a following simple recursive formula. Let $\tau_S(k)$ be the k-th moment of the time required to move from S to E, then

$$\tau_S(k) = \frac{1}{-Q_{S,S}}\left[\tau_S(k-1) + \sum_{x \in E-S} Q_{S,S\cup\{x\}}\tau_{S\cup\{x\}}\right]$$
$$\tau_E(k) = 0$$

for any integer k. Hence, the k-th moment $W_D\left(X_C^G\right)$ is given by $\tau_\phi(k)$, see [1] for details. This formula gives us that the expected value of $W_D\left(X_C^G\right) = 41.025$. This value dominates the maximum premium in the deterministic model by 14%. It is trivial to show that the greater the number of tasks, the larger this percentage is, even if the additional tasks have lower expected values.

Table 3 gives the probability that each basic element in $\mathcal{M}_C$ is the optimal schedule selection. These values were found by computing the probabilities of absorption in the discrete time Markov chain with transition probability matrix P as given in corollary 3. From these absorption probabilities, we compute the criticality indices shown in the bar graph given in Figure 7. In this context, the criticality of a task is the probability that it will be undertaken by a premium maximizing scheduler.

TABLE 3. THE PROBABILITY OF ABSORPTION FOR EACH BASIC ELEMENT IN $\mathcal{M}_C$

$P\left[X = X_c^G\right]$	X
0.109671	{1,2,3,5}
0.158153	{1,2,3,6}
0.049650	{1,2,4,5}
0.096890	{1,2,5,6}
0.053577	{1,3,4,5}
0.076736	{1,3,4,6}
0.103941	{1,3,5,6}
0.080366	{2,3,4,5}
0.115104	{2,3,4,6}
0.155911	{2,3,5,6}

Criticality indices are interesting even for small problems. In the deterministic premium model task four would never be undertaken, yet its criticality is well over 1/3. Also, with deterministic premiums, we would have tasks two, three, and six always scheduled, and these criticalities are between 70% and 86%, far short of certainty.

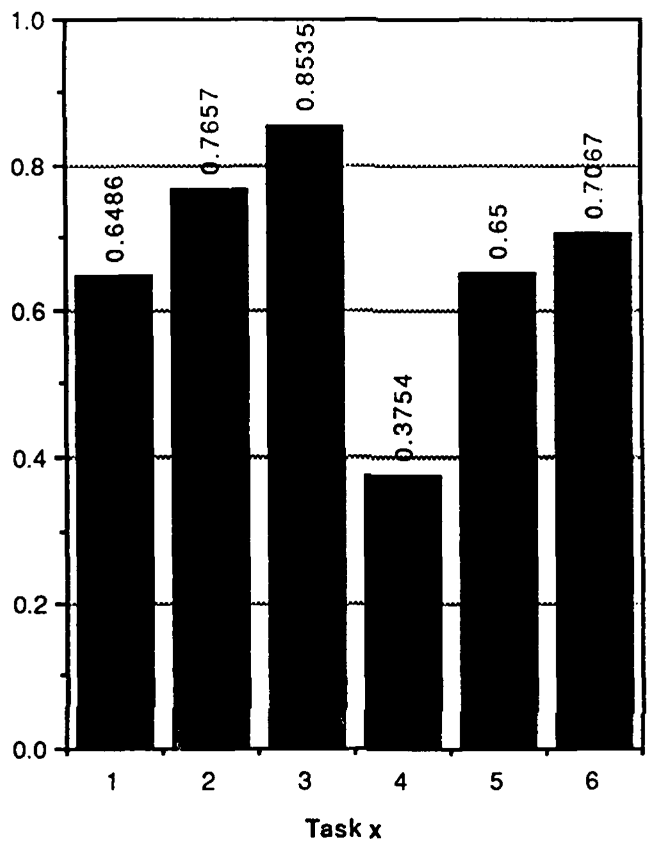

Figure 7. Criticality Indices

5. AN EXAMPLE IN SPANNING TREES

The best known application of matroid optimization is the optimal spanning tree problem. As mentioned above, the minimum weight spanning tree problem has been studied extensively, with the exponentially weighted case investigated in Kulkarni [7].

Let (N,A) be as graph, N being the set of nodes and A being the set of arcs. We wish to find a set of arcs such that there is a unique path from every node to every other. This set is called a spanning tree, and is of obvious interest to communications system designers, among others. Let $|N| = n_c + 1$, so that any spanning tree has n_c arcs. Suppose that each arc has a value v(e), $e \in A$, and that we wish to find the maximum weight spanning tree in the network. The following policy, attributed to Kruskal, see Welsh [9], always finds the maximum weight spanning tree:

> ***POLICY: At each stage, select the highest-valued unselected arc such that the set of selected arcs has no cycles.***

Thus, $\mathcal{M}_C = \{X \subset A:$ there are no cycles in $(N, X)\}$, and $\beta_{\mathcal{M}_C}$ is the set of spanning trees, hence $\beta_{\mathcal{M}}$ is the set of complements to spanning trees. We will illustrate using the network in Figure 8.

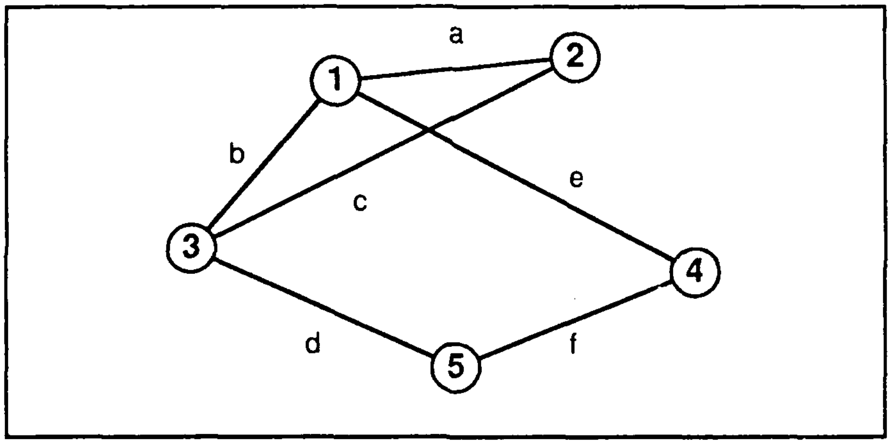

Figure 8. Network for Spanning Trees

TABLE 4. ARCS AND WEIGHTS FOR SPANNING TREES

arc x	E[V(x)]
a	15
b	21
c	13
d	17
e	21
f	19

Using the values given in table 4 as deterministic weights, the algorithm will find the maximum weight spanning tree to be {b, d, e, f}, with a deterministic weight of 76. If the weights are exponentially distributed, the distribution function and density of the weight of the maximum weight spanning tree are given in figures 9 and 10, resp. The expected maximum weight is 107.58, dominating the optimal solution for deterministic weights by 42%.

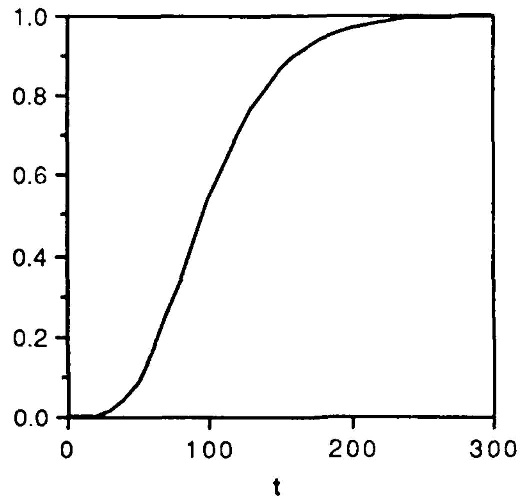

Figure 9. The Distribution of the Weight of the Maximum Weight Spanning Tree

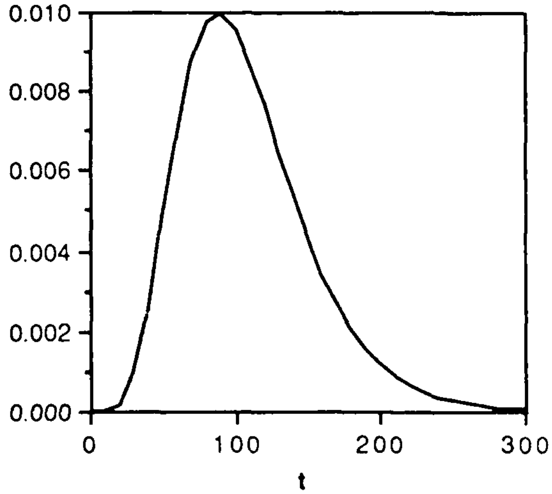

Figure 10. The Density Function of the Weight of the Maximum Weight Spanning Tree

Using the results on the embedded Markov chain, we can calculate the probability that a given basic element is optimal, as well as criticality indices for each of the arcs.

TABLE 5. ABSORPTION PROBABILITIES

$P\left[X = X_c^G\right]$	X
0.12484294	{a,b,d,e}
0.08179365	{a,b,d,f}
0.13953035	{a,b,e,f}
0.06455315	{a,c,d,e}
0.04229344	{a,c,d,f}
0.07214763	{a,d,e,f}
0.08365758	{b,c,d,e}
0.10819722	{b,c,d,e}
0.07088783	{b,c,d,f}
0.12092630	{b,c,e,f}
0.08116990	{c,d,e,f}

These results are given in table 5 and figure 11. Several of the basic elements are nearly equally probable as optimal solutions, with probabilities between 0.10 and

0.14, and none of the basic elements could be said to be extremely unlikely to be optimal. As for the criticality indices, arc e will be a member of the optimal solution 80% of the time, while all of the arcs are at least 50% probable to be included.

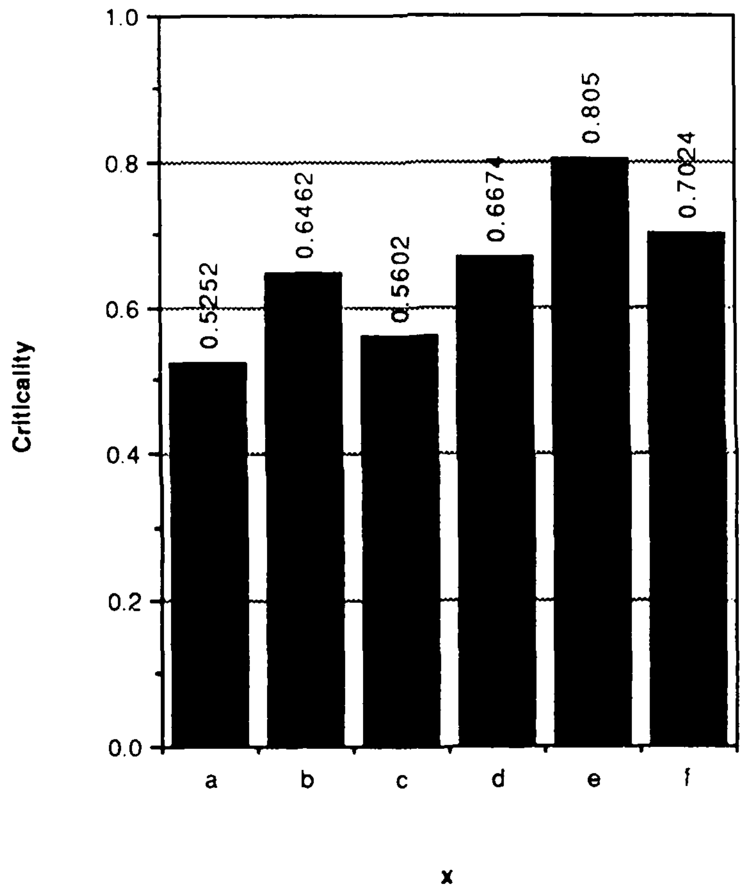

Figure 11. Criticality Indices

6. CONCLUSION

In this work, we have studied the stochastic behavior of the the greedy algorithm on matroids when the ground set element weights are random variables, especially when they are independent exponentials. We showed that, by restructuring the algorithm to a dual minimization algorithm with a discounting mechanism, we could model this new algorithm as a Markov process. This Markov process has absorption times which are distributed identically with the weight of the optimal basic element in the matroid. Using the embedded Markov chain, we are able to calculate the probability that a each basic element is optimal, and we can derive criticality indices for each of the ground set elements.

The results we have provided are applicable to any matroid. To suggest applications, we provided two well known examples of matroid maximization, the transversal matroid used in optimal scheduling and the graphic matroid used to find maximum weight spanning trees. A large portion of the applications of matroid theory are reducible to one of these two examples.

REFERENCES

[1] Bailey, M. P., "Constant Access Systems: A General Framework for Optimization on Stochastic Networks," NPS-55-89-02, Naval Postgraduate School, Monterey, CA 93943-5000.

[2] Bailey, M. P., "Minimization on Stochastic Matroids," NPS-55-90-01, Naval Postgraduate School, Monterey, CA 93943-5000.

[3] Edmonds, J., "Matroids and the Greedy Algorithm," *Mathematical Programming*, **1**, (1971), pp. 127-136.

[4] Edmonds, J. and D. R. Fulkerson. 1970. Bottleneck Extrema. *Journal of Combinatorial Theory* **8**, 299-306.

[5] Frieze, A. M., "One the Value of a Random Minimum Spanning Tree Problem" *Discrete Applied Mathematics*, **10**, (1985), pp. 47-56.

[6] Jain, A. and J. W. Mamer, "Approximations for the Random Minimal Spanning Tree with Application to Network Provisioning," *Operations Research*, **36**, (1988), pp. 585-588.

[7] Kulkarni, V. G., "Minimal Spanning Trees in Undirected Networks with Exponentially Distributed Arc Weights," *Networks* **18** (1988) pp.111-124.

[8] Lawler, E. L., *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart, and Winston, New York (1976).

[9] Welsh, D. J., "Kruskal's Theorem for Matroids," *Proceedings of the Cambridge Philosophical Society*, **64** (1968), p. 3-4.

TABLE 4.2. THE SAMPLE PATH THE ALGORITHM WILL FOLLOW

$X_c \in \beta_{\mathcal{M}_c}$	$X \in \beta_{\mathcal{M}}$	$\omega_D(X_c)$	$\omega(X)$
{1,2,3,5}	{4,6}	31	14
{1,2,3,6}	{4,5}	35	10
{1,2,4,5}	{3,6}	25	20
{1,2,5,6}	{3,4}	29	16
{1,3,4,5}	{2,6}	28	17
{1,3,4,6}	{2,5}	30	15
{1,3,5,6}	{2,4}	32	13
{2,3,4,5}	{1.6}	31	18
{2,3,4,6}	{1,5}	33	12
{2,3,5,6}	{1,4}	35	10

DISTRIBUTION LIST

Library (Code 0142)..2
Naval Postgraduate School
Monterey, CA 93943-5000

Defense Technical Information Center..2
Cameron Station
Alexandria, VA 22314

Office of Research Administration (Code 012) 1
Naval Postgraduate School
Monterey, CA 93943-5000

Center for Naval Analyses...1
4401 Ford Avenue
Alexandria, VA 22302-0268

Library (Code 55)...1
Naval Postgraduate School
Monterey, CA 93943-5000

Operations Research Center, Rm E40-164..1
Massachusetts Institute of Technology
Attn: R. C. Larson and J. F. Shapiro
Cambridge, MA 02139

Koh Peng Hong 1
OA Branch, DSO
Ministry of Defense
Blk 29 Middlesex Road
SINGAPORE 1024

Institute for Defense Analysis..1
1800 North Beauregard
Alexandria, VA 22311

Michael Page Bailey, Code 55Ba...1
Naval Postgraduate School
Monterey, CA 93943-5000

V. G. Kulkarni 1
Department of Operations Research
The University of North Carolina at Chapel Hill
Chapel Hill, NC 27599

J. George Shanthikumar...1
The Management Science Group
School of Business Administration
University of California at Berkeley
Berkeley, Ca 93720

Shaler Stidham..1
Department of Operations Research
CB3180 Smith Building
University of North Carolina
Chapel Hill, NC 27599

Kevin D. Glazebrook..1
Department of Mathmatics and Statistics
University of Newcastle-upon-Tyne
Newcastle-upon-Tyne, NE1 7RN
United KINGDOM

Gideon Weiss..1
School of ISYE
Georgia Tech
Atlanta, GA 30332-0305

David Goldsman...1
School of ISYE
Georgia Tech
Atlanta, GA 30332-0305

David Yao 1
Uris Hall
Columbia University
New York, NY 10027

Rakesh Vohra..1
242 Hagerty Hall
1775 College Rd.
Columbus, OH 43210-1399

USMC Warfighting Center, MCCDC...........................1
Quantico, VA 22134-5001